Journal of Computational Physics 250 (2013) 616–643

Contents lists available at SciVerse ScienceDirect

Journal of Computational Physics

journal homepage: www.elsevier.com/locate/jcp

A nonparametric belief propagation method for uncertainty quantification with applications to flow in random porous media

Peng Chen [a], Nicholas Zabaras [a,b,*]

[a] Materials Process Design and Control Laboratory, Sibley School of Mechanical and Aerospace Engineering, Cornell University, 101 Frank H.T. Rhodes Hall, Ithaca, NY 14853-3801, USA
[b] Center for Applied Mathematics, Cornell University, 657 Frank H.T. Rhodes Hall, Ithaca, NY 14853-3801, USA

ARTICLE INFO

Article history:
Received 10 December 2012
Accepted 8 May 2013
Available online 31 May 2013

Keywords:
Uncertainty quantification
Predictive modeling
Probabilistic graphical models
Porous media flow
Nonparametric belief propagation
Loopy belief propagation

ABSTRACT

A probabilistic graphical model approach to uncertainty quantification for flows in random porous media is introduced. Model reduction techniques are used locally in the graph to represent the random permeability. Then the conditional distribution of the multi-output responses on the low dimensional representation of the permeability field is factorized into a product of local potential functions. An expectation–maximization algorithm is used to learn the nonparametric representation of these potentials using the given input/output data. We develop a nonparametric belief propagation method for uncertainty quantification by employing the loopy belief propagation algorithm. The nonparametric nature of our model is able to capture non-Gaussian features of the response. The proposed framework can be used as a surrogate model to predict the responses for new input realizations as well as our confidence on these predictions. Numerical examples are presented to demonstrate the accuracy and efficiency of the proposed framework for solving uncertainty quantification problems in flows through porous media using stationary and non-stationary permeability fields.

1. Introduction

Uncertainty Quantification (UQ) of multi-scale and multi-physics systems is a field of great interest and has attracted the attention of many researchers and communities in recent years. However, it is difficult to construct a complete probabilistic model of such problems mainly because they often involve high-dimensional and continuous random variables, which have complex, multi-modal distributions.

Over the past few decades, many methods and algorithms have been developed to address UQ problems. The most widely used approach is the Monte Carlo (MC) method. The wide acceptance of the MC method is due to the fact that it can uncover the complete statistics of the solution, while having a convergence rate that is (remarkably) independent of the input dimension. Nevertheless, it quickly becomes inefficient in high dimensional and computationally intensive problems, where only a few samples can be observed. Other methods are attempting to construct a surrogate model for the complex physical system. The idea is to run the deterministic physical solver on a small, well-selected set of inputs and then use these data to learn the

* Corresponding author at: Materials Process Design and Control Laboratory, Sibley School of Mechanical and Aerospace Engineering, Cornell University, 101 Frank H.T. Rhodes Hall, Ithaca, NY 14853-3801, USA. Tel.: +1 607 255 9104.
E-mail addresses: nzabaras@gmail.com, zabaras@cornell.edu (N. Zabaras).
URL: http://mpdc.mae.cornell.edu/ (N. Zabaras).

0021-9991/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jcp.2013.05.006

response surface, so that the UQ problem can be studied based on the surrogate instead of the computationally expensive simulator. Such methods include, adaptive sparse grid collocation method (AGSC) [1], multi-response Gaussian process method (MGP) [2], adaptive locally weighted projection regression methods (ALWPR) [3], and so on. However, all these methods face the "curse of dimensionality", when the inputs are high-dimensional.

Probabilistic graphical models [4] provide a powerful framework that effectively interprets complex probabilistic relations between many inter-correlated variables. The two basic elements of a graphical model are its nodes and edges. The nodes represent the random variables and edges linking nodes represent correlations between them. The joint probability distribution can be accessed by decomposing the complex network into local clusters (e.g., maximal cliques) defined by connected subsets of nodes. Then, by applying appropriate inference algorithms, the marginal and conditional probabilities of interest can be effectively calculated.

Probabilistic graphical model has been used in a range of application domains, which include web search [5], medical and fault diagnosis [6], speech recognition [7], robot navigation [8], bioinformatics [9], communications [10], natural language processing [11], computer vision [12], and many more. Most of these applications involve discrete random variables or low-dimensional continuous random variables. However, for problems involving high-dimensional continuous variables, the number of efficient and accurate algorithms is limited.

The general procedure of studying a graphical model problem can be summarized as follows: (1) Design the structure of the graphical model; (2) Select suitable model reduction techniques to reduce the dimensionality of the random input; (3) Prepare the training data; (4) Learn the parameters of the graphical model using the training data; (5) Solve an inference problem, that is, find the conditional or marginal probabilities of interest. In this work, the problem under consideration is flow through porous (heterogeneous) media. We are interested in constructing a graphical model that captures the probabilistic relationship between the input permeability field and the response properties field, such as velocity and pressure. All the unknown parameters of the graphical model can be learned locally via techniques such as maximum likelihood (MLE), or maximum a posterior probability (MAP). To address the inference problem, several sampling and variational algorithms can be applied. Since the designed graphical model in this work is nonparametric, a sampling-based nonparametric belief propagation [13,14] algorithm is employed to carry out the inference task. The proposed framework allows us to investigate the uncertainty propagation problem. In addition, it can also act as a surrogate model to the deterministic solver, that is, for any realization of the input permeability, it can give us the predictions of physical responses as well as the confidence on these predictions (induced by the limited data used to train the graphical model).

In [15], the authors proposed a probabilistic graphical model for multiscale stochastic partial differential equations (SPDEs) that focuses on the correlation between physical responses. The distribution of physical responses conditioned on stochastic input was approximated using conditional random field theories. Different physical responses (such as flux and pressure in flows in heterogeneous media) are correlated in such a way that their interactions are assumed to be conditioned on fine-scale local properties. No model reduction of fine-scale properties was involved in this process. The influence of fine-scale properties on coarse-scale responses was modeled through a set of hidden variables. The approach in this paper is significantly different in multiple fronts: (1) the graphical model considers output responses that are independent of each other; (2) an explicit model reduction scheme is considered to reduce the dimensionality of the random permeability field without the need for introducing hidden variables; and (3) the graph structure and graph learning scheme are implemented based on the Expectation/Maximization (EM) algorithm and a sampling based approach to nonparametric belief propagation, respectively.

This paper is organized as follows. First, the problem definition is given in Section 2. Then the basic procedure of how to construct an appropriate graphical model and all the associated algorithms are discussed in Sections 3–5. In Section 6, we introduce the porous media flow problem and provide various examples demonstrating the efficiency and accuracy of the graphical model approach. Brief discussion and conclusions are finally provided in Section 7.

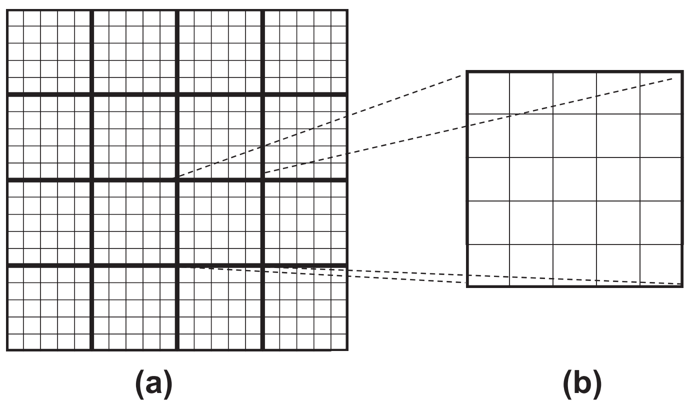

Fig. 1. Schematic of the domain partition: (a) fine- and coarse-scale grids and (b) fine-scale local region in one coarse-element.